Renewable and Sustainable Energy Reviews 20 (2013) 411–419

Contents lists available at SciVerse ScienceDirect

Renewable and Sustainable Energy Reviews

journal homepage: www.elsevier.com/locate/rser

# A review of sensitivity analysis methods in building energy analysis

Wei Tian*

*Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK*

ARTICLE INFO

*Article history:*
Received 20 October 2012
Accepted 16 December 2012
Available online 11 January 2013

*Keywords:*
Building energy
Sensitivity analysis
Global method
Variance-based method
Correlated inputs

ABSTRACT

Sensitivity analysis plays an important role in building energy analysis. It can be used to identify the key variables affecting building thermal performance from both energy simulation models and observational study. This paper is focused on the application of sensitivity analysis in the field of building performance analysis. First, the typical steps of implementation of sensitivity analysis in building analysis are described. A number of practical issues in applying sensitivity analysis are also discussed, such as the determination of input variations, the choice of building energy programs, how to reduce computational time for energy models. Second, the sensitivity analysis methods used in building performance analysis are reviewed. These methods can be categorized into local and global sensitivity analysis. The global methods can be further divided into four approaches: regression, screening-based, variance-based, and meta-model sensitivity analysis. Recent research has been concentrated on global methods because they can explore the whole input space and most of them allow the self-verification, i.e., how much variance of the model output (building energy consumption) has been explained by the method used in the analysis. Third, we discuss several important topics, which are often overlooked in the domain of building performance analysis. These topics include the application of sensitivity analysis in observational study, how to deal with correlated inputs, the computation of the variations of sensitivity index, and the software issues. Lastly, the practical guidance is given based on the advantages and disadvantaged of different sensitivity analysis methods in assessing building thermal performance. The recommendations for further research in the future are made to provide more robust analysis in assessing building energy performance.



**Contents**



* Tel.: +44 1223 748595.
*E-mail address:* tjtianjin@yahoo.com.cn

1364-0321/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.rser.2012.12.014

## 1. Introduction

Sensitivity analysis is a valuable tool for both energy simulation models and observational study in building energy analysis. Therefore, sensitivity analysis has been widely used to explore the characteristics of building thermal performance in various types of applications, such as building design [1,2], calibration of energy models [3,4], building retrofit [5,6], building stock [7,8], impact of climate change on buildings [9,10].

The methodology for sensitivity analysis is the same in different types of application in building energy analysis. Fig. 1 shows typical steps for implementing sensitivity analysis in building performance analysis: determine input variations; create building energy models; run energy models; collect simulation results; run sensitivity analysis; presentation of sensitivity analysis results. The main difference among these applications is the variations (uncertainty or probability distributions) of input factors for different research purposes, which is often overlooked or not emphasized enough in the field of building analysis. For instance, the thermo-physical properties of building envelope can be regarded as normal distributions when comparing simulated energy to actual monitored energy consumption [11]. This is because the variations of inputs in this case are due to natural variations because of construction, aging, and actual conditions of buildings. However, they should be defined as uniform distributions when the aim is to identify effective energy saving measures in building design [12]. This is because the design variables can be regarded as being equally probable. It is also possible to combine these natural uncertainties (normal distributions) and the design ranges (uniform distributions) together for building physical properties. With this end in view, more complicated methods are needed, such as two-dimensional Monte Carlo analysis [13].

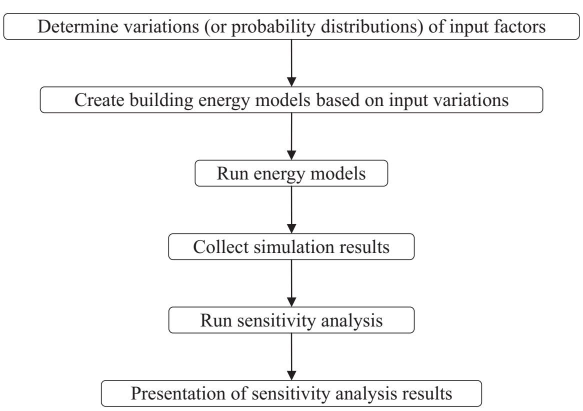


**Fig. 1.** Typical schematic flow diagram for sensitivity analysis in building performance analysis

The methods for sensitivity analysis applied in the domain of building analysis can be divided into local and global approaches [14] as summarized in Table 1. Local sensitivity analysis is focused on the effects of uncertain inputs around a point (or base case), whereas global sensitivity analysis is more interested in the influences of uncertain inputs over the whole input space [15]. Therefore, the global approach is regarded as a more reliable method. The disadvantages of using global methods are high computationally demanding compared to local sensitivity analysis. Both local [6,16–24] and global methods have been widely used in building performance analysis. The global sensitivity analysis includes regression methods[1,2,7–9,13,25–34], screening-based [35–41], variance-based [12,15,42–44] and meta-modelling approaches [9,45,46]. The characteristics of these methods have been summarized in Table 1.

This paper is to review the application of sensitivity analysis in building thermal performance analysis, provide the practical advice on how to effectively use sensitivity analysis in different settings, and make recommendations for applying sensitivity analysis in building performance analysis in the future. This paper is structured as follows. Section 2 describes the typical steps for implementing sensitivity analysis in building performance analysis. Section 3 reviews the methods used in the field of building energy analysis and also comments on the advantages and shortcoming for these sensitivity analysis approaches. Section 4 focuses on several topics which are important, but often overlooked when implementing sensitivity analysis in building analysis. These topics include how to deal with correlated inputs, calculate the variations of sensitivity measures, the differences between observational data and simulation models in using sensitivity analysis.

**Table 1**
Comparison of sensitivity analysis methods used in building performance analysis

| | Method | Subtype | Characteristics | Literature |
|---|---|---|---|---|
| Local | Local | – | Explore a reduced space of the input factor around a base case; low computational cost; simple to implement; easy to interpret; not consider interactions between inputs; no self-verification | [6,16–24] |
| Global | Regression | SRC | SRC and $t$-value, suitable for linear models; SRRC, suitable for non-linear but monotonic models; | [1,7–9,25–27,29,34] |
| | | SRRC | moderate computational cost for energy models; fast to compute; easy to implement and understand; | [2,13,25,30–32] |
| | | $t$-value | high SRC means more important of the variable | [10,33] |
| | Screen | Morris | Suitable for a larger number of inputs and computationally intensive models; model-free approach; qualitative measure to rank factors; no self-verification; not suitable for uncertainty analysis | [35–41] |
| | Variance based | FAST | Decompose the variance of the model output for every input; model-free approach; consider both main and | [12,42] |
| | | Sobol | interactions effects; quantitative measures; high computational cost; FAST is not suitable for discrete distributions | [44] |
| | Meta-model | MARS | Suitable for complex and computationally intensive models; | [7,45] |
| | | ACOSSO | quantify output variance due to different inputs; | [9] |
| | | SVM | the accuracy dependent on the meta-model | [46] |

*Notes*: SRC, standardised regression coefficients [47]; SRRC, standardized rank regression coefficient [47]; FAST, Fourier amplitude sensitivity test [49]; MARS, multivariate adaptive regression splines [52], ACOSSO, adaptive component selection and smoothing operator [52], SVM, support vector machine [61].

## 2. Typical steps for sensitivity analysis in building performance analysis

This section will describe the typical steps for applying sensitivity analysis in building performance analysis as shown in Fig. 1. We also discuss some practical issues, such as the determination of input variations, the choice of building simulation programs, how to speed up the calculation for building energy models.

### 2.1. Input variations

The first step in sensitivity analysis is to determine the range of the inputs. The probability distributions of the inputs also need to be defined if using sampling-based methods, such as variance-based or Monte Carlo sensitivity analysis [47]. The ranges (or distributions) of inputs are mainly dependent on the research purposes for sensitivity analysis. The following discussion will be based on three different settings: assess the energy performance in a new building using different design options; explore the variations of energy use in an existing building; perform the retrofit analysis for an existing building using different energy saving measures.

The first example is to compare the energy performance for a new building using different design options. The analyst should focus on the possible ranges for the design variables. In this case, the input variables can be taken as discrete or continuous uniform distributions if we assume that these variables are equally probable [12]. For example, the *U*-values for building envelope depend on different types of insulation material or different insulation thickness and, consequently, the *U* values are equally likely to happen in the design phase. Then sensitivity analysis can be used to decide whether higher insulation is necessary compared to other energy saving measures.

The aim of the second example is to explore the possible variations of energy use for an existing building in actual use and find out the key variables affecting them [11]. Another purpose for this example may compare the simulated and measured energy consumption and determine the main reasons to explain the difference between them. Then the analyst should concentrate on the possible ranges for building input factors in a specific building. For example, the *U*-value for building envelope may be regarded as normal distributions in this case [9]. This is because the *U*-value is likely to be (or at least approximately) a constant and the variations of *U*-values may be due to insulation quality, construction, temperature, age, maintenance etc [48]. Hence, normal distributions may be a good choice for most of the variables in the analysis. The variables themselves may be a constant value for an existing building and the variations are because of natural degradation, lack of knowledge, or other reasons. These distributions may need to be truncated to avoid negative values in some cases [13].

The third example is related to the optimization of an existing building by using different insulation thickness and other measures. In this case, more complicated methods might be needed. This is due to the combination of the natural variations and the design variations of *U*-values for building envelope. Two-dimensional Monte Carlo method can be used in this case to identify the key variables by considering these two types of variations of input factors [13]. This analysis includes two loops to include the two types of effects separately. Hence, it is computationally expensive and the results are hard to interpret even for a professional analyst. A simple solution is to only consider the design variations because the effects of different design options on energy performance may be more significant than those from natural variations in buildings.

It should be noted that how to define the input ranges would also depend on the choice of sensitivity analysis method. For local sensitivity analysis, it is very straightforward since it does not need sampling methods to generate combinations of inputs. For most of global sensitivity analysis, it is necessary to implement sampling strategies for generation of sample. For regression and meta-model methods, Latin-hypercube sampling (LHS) is very popular due to its efficient stratification properties [47]. LHS has been widely used in building performance analysis [2,7,25,28,29,43,45]. For screening and variance-based methods, they usually require special sampling methods for these sensitivity analysis approaches [49].

### 2.2. Create building energy models

For most of sensitivity analysis methods, it involves a large number of simulation runs. Hence, it is necessary to automate the process of creating building energy models with the different combinations of the inputs from the last step. A number of building simulation programs have been used in sensitivity analysis, such as EnergyPlus [1,9,22,25,31,46,50], ESP-r [11,35], TRNSYS [27,43], and DOE2 [17]. These programs are very flexible to be suitable for sensitivity analysis. For example, the model input files for both EnergyPlus and DOE2 are text files, which can be easily processed using computer languages, such as Excel VBA [25], Matlab [43]. Another popular choice is based on simplified energy models [33,36,37], which are fast to compute and easy to change inputs. These energy calculation methods are often based on ISO 13790:2008 (Thermal Performance of Buildings – Calculation of Energy Use for Space Heating and Cooling) or other simple steady-state energy balance approaches.

### 2.3. Run building energy models

This step is to run a number of simulation models created from building energy simulation programs. This step is usually the most time-consuming process in the sensitivity analysis in terms of the computer time. The calculation for sensitivity analysis usually involves many independent jobs, i.e., many separate building simulation models. Hence, the parallel computing can be very helpful to speed up the calculation. There are two types of methods for parallel computing: single computer with multi-core or multi-processor; multiple computers.

Modern computers usually have multi cores or processors, which can be used for this purpose. Some building energy simulation programs have capability to use this feature. For example, EnergyPlus program can allow better utilization of the computer power by specifying the number of simultaneous simulation processes [51].

Another way to expedite the calculation is using multiple computers. Tian et al. [9] used a campus Condor grid (PlymGrid, around 200 computers) to run 2400 EnergyPlus models in order to investigate the influences of climate change on energy performance for a campus building. It took only one day to finish all these simulation runs using Plymgrid, while it would need around one month if using one normal specification office computer. Hygh et al. [1] used 11-node, 88-core Linux cluster to run 20000 EnergyPlus models to quantify the sensitivity of heating, cooling, and total energy loads in four locations, USA [1]. Eisenhower et al. [46] implemented a 184-CPU Linux cluster to run 5000 EnergyPlus models.

### 2.4. Collect simulation results

This step is very straightforward to collect the results from building energy simulation. For sampling-based methods, it

usually involves a lot of data and script languages can be used to automate this process.

### 2.5. Run sensitivity analysis

This step is to run sensitivity analysis based on the inputs and outputs obtained from the earlier steps. The sensitivity analysis itself is usually not time consuming compared to running building energy models. However, this step may take several minutes or even longer if using meta-model methods with larger number of input factors [52]. This is because some types of meta-models are slow to build although it takes much less time in comparison with running full detailed building energy simulation models. After constructing meta-models, it is fast to run them for sensitivity analysis.

### 2.6. Presentation of sensitivity analysis

There are many ways to present the results from sensitivity analysis, such as scatter plot, Tornado plot, box plot, spider plot. Scatter plots are a good starting point to explore the relationship between inputs and outputs [47]. Burhenne et al. [53] applied scatter plots to investigate the effects of building parameters on energy performance in a typical German building. Tornado plots are very useful to compare the relative importance of different input factors [1,2,25,28]. The largest bar appears at the top to the plot and the second largest appears after that, and so forth. Pie chart is also very popular to show the relative importance of variables in building energy analysis [12,43].

Morris proposed a graphical representation to show the importance of each input based on Morris method, which will be described in Section 3.2.2 [54]. The mean and standard deviation of the elementary effects for different inputs from Morris method are plotted in a two-dimensional graph. Therefore, the low values of both the mean and standard deviation (in the lower left corner of the plot) indicate a non-influential input, while the inputs in the upper right corner of the plot are key variables. The Morris plots have been used in many studies in building analysis [35,37–40].

Spider and box plots are not widely used for sensitivity analysis in the field of building performance analysis. The reason for this is that the sensitivity analysis methods related to spider and box plots are rarely used in assessing building energy performance. These two types of plots do have some advantages to show the results from sensitivity analysis. Spider plots are used to show the impact of variations of input variables on outputs in detail [55]. The advantage of using spider plots is that it can show whether there are linear or non-linear relationship between inputs and outputs. Box plots are mainly used to show the variations of sensitivity measures for different input factors [52]. However, most of studies in building performance analysis did not calculate the variations of sensitivity index.

## 3. Sensitivity analysis methods used in building performance analysis

The section will review the application of sensitivity analysis in building energy analysis. For the technical details of these methods, please refer to [47,49,52,56]. The focus in this section is on the advantages, disadvantages, applicability, and examples of these methods in assessing building thermal performance. Table 1 summarizes the characteristics of these methods and the literature in which these sensitivity analysis methods have been used to assess building thermal performance.

### 3.1. Local sensitivity analysis

Local sensitivity analysis (also called differential sensitivity analysis) belongs to the class of the one-factor-at-a-time methods. Sensitivity measures are usually calculated when one factor is changed and all other factors are fixed. As a result, the choice of base case is very important in this method. This method has apparent advantages. It is very straightforward compared to global sensitivity analysis. As a result, it is easily applied and interpreted. It often needs less simulation runs in comparison with global sensitivity analysis. However, this method has the following drawbacks [57,58]. First, it only explores a reduced space of the input factor around a base case. Second, the interactions cannot be considered using this method. Third, there is no self-verification in this method, while most of global sensitivity analysis can be used to explain how much variations of the outputs are accounted by the input factors [57].

This local sensitivity analysis has been used extensively in the field of building energy analysis [6,16–24]. Rasouli et al. [18] applied local sensitivity analysis to explore the thermal performance for a two-storey office building in Chicago, Illinois, USA The results indicate that the most important factor for HVAC system energy is the ventilation rate. Lam et al. [17] investigated the energy performance of an office building in Hong Kong using DOE-2 program. The results indicate that annual energy and peak design loads are more sensitive to internal loads, window system, temperature set-points, and HVAC equipment efficiency. Demanuele et al. [22] used differential sensitivity analysis to determine the key factors affecting the total energy use in a UK school. It is found that the important variables are related to occupants, such as office and class equipment load and hours of use, heating schedule and set-point temperatures.

### 3.2. Global sensitivity analysis

#### 3.2.1. Regression method

Regression method is the most widely used method for sensitivity analysis in building energy analysis [1,2,7–9,13,25–34]. This is because this method is fast to compute and easy to understand. Many indicators can be used for this purpose usually after Monte Carlo is being performed, such as SRC (Standardised Regression Coefficients), PCC (Partial Correlation Coefficients), and their rank transformation (SRRC standardized rank regression coefficient, PRCC partial rank correlation coefficient). SRC and PCC are only suitable for linear models and the rank transformation (SRRC and PCC) can be used for non-linear but monotonic functions among inputs and outputs. If there is no correlation in inputs, SRC and PRC give the same results for ranking the importance of inputs. The difference between SRC and PRC is that PRC is suitable for correlated input because it excludes the effects of correlations between input factors, but the SRC is only valid in the case of uncorrelated inputs [47]. However, there are many new sensitivity analysis methods to deal with correlated inputs, which will be described in Section 4.2.

Another commonly used technique in regression method is stepwise. The forward stepwise method is the most widely used in sensitivity analysis [47]. The most importance factor firstly enters the model and then the next important factor also enter the model, and repeat this process until no variables are significant in terms of statistical test. The selection criterion includes SRC, $t$-values, adjusted $R$-square, and Akaike information criterion.

The sensitivity index SRC has been widely used in building energy analysis [1,7–9,25–27,29,34]. Domínguez-Muñoz et al. [26] applied SRC method to determine the key variables influencing the peak cooling load of a perimeter zone (intermediate floor) in a three-story office in the south of Spain. The results

indicate that the two most important factors are internal thermal mass and convective heat transfer coefficient between the internal mass and air. The remaining uncertainties are due to solar gains, internal heat gains, and actual ventilation level. Hygh et al. [1] used EnergyPlus as a simulation tool to explore the energy performance of office buildings in four USA cities by using SRC sensitivity indicator. The results show that the influences of design parameters on building energy use vary by different climate zones. Ballarini et al. [29] used SRC to determine the key variables affecting the cooling energy for a residential building in Italy. The results show that the first three most important factors are solar shading, window area, and window insulation. Breesch et al. [27] applied SRC to determine the most influential factors affecting thermal comfort with cross night ventilation on the south side of a typical office building in Belgium. They found that internal heat gains and air tightness are the two most important variables.

The use of sensitivity index SRRC is also reported in building performance analysis [2,13,25,30–32]. Yildiz et al. [25] investigated annual cooling load in low-rise apartment in hot-humid climate using SRRC method with EnergyPlus program. They found that the most sensitive parameters affecting annual cooling energy are natural ventilation, window area, and solar heat gain coefficient (SHGC). De Wilde et al. [13] used SRRC as sensitivity measure to determine the major contributors for heating energy use in a mixed-mode office building in the UK The results indicate the infiltration rate, lighting gains, and equipment gains are the three most important factors in this building.

Many other statistics can be also used to determine which factors are important in regression analysis. These statistics include $t$-value, $F$-value, change of $R^2$ (coefficient of determinations). The $t$-value is the statistic used to test whether the coefficient of the corresponding variable is zero. The higher the absolute value of $t$ (or $F$, change of R2), the more important is the corresponding variable. Zhao [33] used $t$-value with stepwise method in building stock analysis of USA They found that the power use intensity from equipment is the most important factor for primary energy use intensity in Chicago area. De Wilde et al. [10] also used $t$-value to show that internal heat gains are important factors for both heating energy and overheating risk in an office building in the UK

#### 3.2.2. Screening-based method

The purpose of screening method is often to fix some input factors from a large number of factors without reducing the output variance [56]. The Morris method is the mostly used screening method in the field of building performance analysis [35–41]. The Morris method belongs to global sensitivity analysis because the baseline changes in every step and the final sensitivity measures are calculated by averaging at different points of the input space. Input factors are taken as a discrete number of values (also called levels), which are different from other global methods in which input values are directly from distributions. Two sensitivity indexes can be obtained from Morris method [56]. One ($\mu$) is to estimate the main effect of the input factor on the output and the other ($\sigma$) is to assess the interaction with other factor or the non-linear effects. A new measure ($\mu^*$) has been proposed to estimate the total effects of the input factor [56]. This method is more suitable when there are a few influential factors and a majority of non-influential factors in the project. The main advantage using this method is low computation cost compared to other global sensitivity analysis. For example, the total simulation runs is only 36 if there are 8 inputs and 4 levels for every input [59]. The drawback of this method is that this approach tends to provide qualitative measures by ranking input factors, but it cannot quantify the effects of different factors on outputs. As a result, this method does not allow self-verification, which means the analyst does not know how much of the total variances of outputs have been taken into account in the analysis. The other types of global sensitivity analysis (such as regression or variance based methods) can usually provide this information. Another disadvantage for Morris method is that this method does not provide the uncertainty analysis for building energy performance because the sampling method used in this method cannot converge to the population mean of the model output [59].

The Morris method has been widely used in building energy analysis [35–41]. Heiselberg et al. [38] implemented Morris method to identify the key variables influencing energy use for an office building in Denmark. The results show that the two most important variables are lighting control and ventilation during the winter. Hyun et al. [41] used Morris method to investigate the performance of natural ventilation in a 15-story residential building in Korea. It is found that the four important factors are wind velocity, window opening area by occupants, local terrain constant, and flow exponent. Corrado et al. [37] used Morris method to identify the key factors for energy rating in a house in Italy. They found that the five key factors in decreasing order of importance are indoor temperature, air change rate, number of occupants, metabolic rate, and equipment heat gains.

#### 3.2.3. Variance-based method

The variance-based method is to decompose the uncertainty of outputs for the corresponding inputs [56]. Two main sensitivity measures used in this approach are the first order and total effects. The first order effects consider the main effects for the output variations due to the corresponding input. The total effects account for the total contributions to the output variance due to the corresponding input, which include both first order and higher-order effects because of interactions among inputs. Hence, the difference between the first order and total effects can show the effects of interactions between variables. If the objective of the research is to fix the factors which are not important in the energy models, the total sensitivity effects should be used. In contrast, if the purpose is to prioritize energy saving measures, the first order effects are a better choice.

This variance-based method is regarded as model free approach, which means it is suitable for complex nonlinear and non-additive models. This method can quantify all the variance of the output due to every input and it can also consider the interaction effects among variables. The drawback of this approach is its high computational cost. Two commonly used methods are FAST (Fourier Amplitude Sensitivity Test) and Sobol [14]. The classical FAST method considers only the nonlinear effects, but not interaction effects. The Sobol method can decompose all the output variance, which means no variance for the output is left in the analysis. However, the Sobol method is much more computational expensive compared to other global sensitivity analysis methods. For instance, the total number of simulation runs is 608 if there are 8 input factors and the second order effects are needed in the analysis [59].

Both FAST and Sobol methods have been used in exploring building energy performance [12,15,42–44]. Mechri et al. [12] implemented FAST method to identify the key design variables affecting building thermal performance for a typical office building in Italy. The results indicate that the envelope transparent surface ratio is the most important factor for both heating and cooling energy. Spitz et al. [44] used the Sobol method with 6669 simulation runs to identify the most influential parameters for an experimental house in France. The six important factors affecting the air temperature are: heating capacity, infiltration, fibreglass

thickness, heat exchanger efficiency, internal heat gains, and fibreglass conductivity. Ruiz et al. [43] implemented variance-based method to determine the key factors affecting final energy consumption in office buildings in six Europe cities. The results show that the first order effects can account more than 90% of the variance of outputs and the climate is the most influential factor influencing final energy consumption.

#### 3.2.4. Meta-model based method

Meta-modelling sensitivity analysis is a two-stage approach [14]. First, a meta-model is created using non-parametric regression methods, which do not have a predetermined form (such as linear or nonlinear regression) and consequently it can be suitable for complex models. Second, sensitivity measures are calculated using this meta-model based on variance-based method. The meta-model is to approximate the objective functions using statistical (or machine learning) models [60]. The main point of using meta-models is that running meta-models needs much less time than running detailed building energy simulation models. Hence, this meta-modelling method can provide more efficient sensitivity index compared to variance-based method [49]. This method can also quantify the variance of the output for different input factors since it used variance-based method as described in Section 3.2.3.

The commonly used meta-models in sensitivity analysis include MARS (Multivariate Adaptive Regression Splines) [52], ACOSSO (Adaptive COmponent Selection and Smoothing Operator) [52], Support vector machine [61], GP (Gaussian process) [62] and TGP (treed Gaussian process) [55]. MARS combines spline regression, stepwise model fitting, and recursive partitioning [52]. ACOSSO can be regarded as a weighted or adapted COSSO (component selection shrinkage operator) which uses a rescaling norm to allow for more flexible estimation of the functional components [52]. GP model generalized multivariate Gaussian distributions over finite dimensional vectors to infinite dimensionality and has become very popular in the machine-learning field [62]. Support vector machine (SVT) is a very popular choice in supervised learning, which produces nonlinear boundaries by constructing a linear boundary in a large and transformed version of the feature space [61]. TGP is to provide fully Bayesian non-stationary and nonlinear regression models in order to improve prediction power [55]. For detailed descriptions on these models, please refer the corresponding documents listed here.

This method is still new in the field of building performance analysis [9,45,46]. De Wilde et al. [45] used MARS meta-model to assess the influences of climate change on a UK office building. They found that the three important variables influencing overheating risk in this building are lighting gains, equipment gains, and weather conditions, which can account for more than 90% of variations of outputs. Tian et al. [9] implemented ACOSSO method to investigate the thermal performance of a campus building in the UK The results show that the four key variables affecting annual carbon emissions are lighting gains, solar heat gain coefficients of windows, cooling degree days, and equipment heat gains. These four factors are responsible for around 95% of the output variances. Eisenhower et al. [46] used a meta-model method for sensitivity analysis based on support vector machine to assess energy performance in a two-storey multi-functional building, which includes a drill deck, offices, and administrative rooms.

## 4. Other topics related to sensitivity analysis in building performance analysis

This section will discuss four topics which are related to the application of sensitivity analysis in assessing building thermal performance: apply sensitivity analysis in observational study; deal with correlated input factors; calculate the variations of sensitivity index; choose software for sensitivity analysis.

### 4.1. Observational study

The discussion so far is only related to sensitivity analysis from building energy simulation model in which the design of computer experiments can be used with classic design experiments or modern sampling methods [63]. Sensitivity analysis can be also used in observational study in the field of building energy analysis. Since the input data in this case is directly from actual buildings, the sampling method cannot usually be implemented. As a result, the following methods are not valid in this case: local, screening and variance-based method. Regression and meta-model method for sensitivity analysis are suitable for this application.

The studies on observational studies using sensitivity analysis are mostly related to building stock in the domain of building energy analysis. The analyst usually obtains the inputs and outputs from building survey in a number of buildings. Then sensitivity analysis can be used to determine the key factors affecting building performance (often energy use intensity standardized by floor area) in these buildings. Moran et al. [8] applied step-wise regression to obtain energy models for historic dwellings in Bath, UK The statistical analysis from SRC indicates that the number of open flues is the most important factor influencing heating energy consumption in these buildings. For electricity use, the two important factors are the number of occupants and the use of an electric range. Zhao et al. [34] investigated energy performance of commercial and hotel buildings using stepwise regression and SRC in China. The results indicate that building age is the only significant factor for commercial buildings, while the three important factors for hotel buildings are building area, total number of guest room, and equivalent number of guest rooms.

It should be noted that the inputs from building stock may be correlated. For example, building age may be correlated with building area, which means that the buildings with higher or smaller area are likely built for a specific time period. Another example is that building geometry parameters may be correlated in a given building stock. For instance, building floor number may increase with building floor area to some extent in a building stock. Hence, it is necessary to take into account the effects of these correlated inputs in applying sensitivity analysis in building stock research. However, most of previous research did not consider them. The next section (Section 4.2) will focus on this topic.

### 4.2. Correlated inputs

It is not uncommon that input parameters may be correlated in building energy analysis as described in the end of the last section (Section 4.1). When the inputs (also called predictors) are correlated, this phenomenon is called collinearity (or multicollinearity). The collinearity leads to large variances of some estimated regression coefficients, and then may lead to unstable regression equations [64]. Therefore, the regression equation is not easily interpretable as the regression models from uncorrelated inputs. The SRC or *t*-values as discussed in Section 3.2 cannot be used in the presence of correlated factors.

Many methods have been developed to deal with the correlation inputs for assessing relative importance of inputs [65–67]. Here, we recommend three new methods: LMG (Lindeman, Merenda, and Gold) [66], PMVD (Proportional marginal variance decomposition) [66], and conditional variable importance from random forest [68]. Both the LMG and PMVD methods can

**Table 2**
*R* packages for sensitivity analysis

| Package | Sensitivity analysis method | Characteristics |
|---|---|---|
| *R* sensitivity [70] | SRC, SRRC, PCC, PRCC, Morris, FAST, Sobol | Confidence intervals of sensitivity analysis |
| *R* CompModSA [73] | SRC, SRRC, MARS, ACOSSO, GP | Suitable for complex models (at least 50 simulation runs), but for uncorrelated inputs; bootstrapping confidence intervals |
| *R* tgp [55] | TGP | Suitable for complex model, but uncorrelated inputs |
| *R* relaimpo [71] | LMG, PMVD | Suitable for correlation inputs from observational data, but assume linear models |
| *R* party [74] | Conditional permutation importance from random forest | Suitable for correlated inputs from non-linear models |

*Notes*: SRC, standardised regression coefficients [47]; SRRC, standardized rank regression coefficient [47]; PCC, partial correlation coefficients [47]; PRCC, partial rank correlation coefficient [47]; FAST, Fourier amplitude sensitivity test [49]; MARS, multivariate adaptive regression splines [52], ACOSSO, adaptive component selection and smoothing operator [52], GP, Gaussian process [62]; TGP, treed Gaussian process [55]; LMG, Lindeman, Merenda, and Gold [66]; PMVD, Proportional marginal variance decomposition [66].

decompose the coefficients of determination ($R^2$) into non-negative values by averaging over orderings. The PMVD method is closer to a conditional index in comparison with the LMG approach. These two approaches are only suitable for linear models. If the results from LMG or PMVD method indicate that there is a large variation of the outputs unexplained (i.e., non-linear effects in the model), the conditional variable importance from random forest can be used [68]. This new index is based on random forest approach, which is a popular machine learning algorithm to handle non-linear models and large databases. Strobl et al. [68] proposed this new conditional variable importance (similar to partial correlation) to be more suitable in the case of correlated factors. Note that all these three methods are computationally expensive, especially for a large number of input factors or simulation runs. For more detailed discussion on these three methods and other relative importance approaches suitable for correlated inputs, please refer to [65–68].

For building energy simulation models, the analyst can also consider the correlation structure of inputs in the generation of sampling as long as the analyst has these correlated data or correlation matrix among input factors. For example, the actual lighting power density may be correlated with the occupancy density in a building. Three methods can be used for this case: Iman Conover, Dependency tree, and Stein [56]. The first method (Iman Conover) can directly induce a rank correlation among inputs. The second method (Dependency tree) can define the correlation structure by using undirected acyclic graphs to form a tree structure. The third method (Stein) can use the correlated sample in generating the new sample for inputs. Please refer to [56,59] for detailed descriptions on these methods.

### 4.3. Variations of sensitivity index

Sensitivity indexes from most of previous studies are only point estimate in the field of building energy analysis. Compared to point values, the intervals for sensitivity indexes can provide robust results from sensitivity analysis. Bootstrapping is a common method used to create the variations of different sensitivity indexes. It the data can be assumed as an independent population, a new data set can be obtained by random sampling with replacement from the original dataset [69]. This process is repeated until all the samples required are created in this way.

The bootstrapping method has been implemented in several packages to construct the intervals for sensitivity index. *R* sensitivity package [70] has this function to calculate the intervals for various sensitivity index, such as SRC, PCC, SRCC, PRCC Another *R* package relaimpo package [71] also can calculate the intervals for sensitivity analysis, especially in the case of correlated factors. For meta-model sensitivity analysis, Storlie et al. [52] has developed a *R* package to compute the confidence intervals for sensitivity measures using bootstrapping method.

### 4.4. Software issues

There are many statistical programs for sensitivity analysis. We recommend two programs: Simlab [59] and *R* [72]. Both of them are free and have many types of sensitivity analysis approaches described in this paper.

Simlab is a free statistical program for uncertainty and sensitivity analysis. Many researchers have used this program in assessing variable importance in building performance analysis [2,9,12,13,28,37,40,43]. Simlab has several different types of sensitivity analysis methods, which include SRC, SRRC, Morris, Sobol, FAST Simlab has two versions: a legacy version with a standard desktop application with graphical user interface; a new version with Matlab, Fortran or C/C++ compiler. These features make Simlab very flexible to be suitable for different applications.

*R* is an open-source free software environment, which has cutting edge methods developed by leading statisticians around the world [72]. Another advantage of using *R* is to promote reproducible research by using *R* and sharing *R* code in the field of building performance analysis. *R* has many different types of sensitivity analysis methods with *R* packages as listed in Table 2. *R* sensitivity package [70] has similar functions to Simlab and it also can calculate intervals of sensitivity index as described in Section 4.3. *R* CompModSA package [73] can deal with non-linear models using non-parametric regression methods. *R* tgp package [55] has used treed Gaussian process models for sensitivity analysis in which a unique feature is to show how the output responds to every input variable by considering the variations of main effects. *R* relaimpo [71] and party [74] packages can handle the correlated inputs to assess the relative importance of variables. The difference is that the former (*R* relaimpo) is only suitable of linear models, while the latter (*R* party) can be used for non-linear models.

## 5. Conclusions and further work

This paper has reviewed the methods of sensitivity analysis in building energy analysis. The findings and recommendations of the application of sensitivity analysis in assessing thermal performance in buildings are summarized as follows.

(1) The local sensitivity analysis is the simplest method and still very useful in building performance analysis even with its shortcomings. This is due to its low computational cost, simple implementation, and easy interpretation. The limitation of this

method is that it only explores small portion of the possible space of input values.

(2) The global sensitivity analysis methods have been increasingly applied in building energy analysis to identify the key variables affecting building thermal performance. It is necessary to automate the process of creating energy models and collecting the results from simulation because of more simulation runs for global sensitivity analysis. Hence, the building simulation programs with only GUI (graphical user interface) are hard to implement this method.

(3) The choice of sensitivity analysis methods depends on many factors, which include the research purpose, the computational cost of energy models, the number of input variables, the analyst's time for a project, the familiarity of sensitivity methods. The practical advice of using sensitivity analysis is as follows. (a) The regression method may be still the first choice because it only requires moderate computational cost in the field of building performance simulation. If the results from regression method indicate there is a large proportion of the output variance unexplained by regression models, the meta-model sensitivity analysis can be used to determine the most influential factors in the analysis without running extra energy simulation runs. (b) For high computational energy models, a good choice can be the Morris method if there are a large number of input variables and the analyst only needs qualitative analysis, whereas a better choice may be the meta-model sensitivity analysis in order to quantify the variance of output for every input. (c) The variance-based method can give more reliable results at the cost of increased computational time. (d) For observational study on building energy performance, the regression or meta-model sensitivity analysis can be used, while the screening or variance-based method is inappropriate. This is because the screening or variance-based methods are dependent on the sampling approaches for the input factors.

(4) The variations of inputs in the analysis depend on the research purpose. If the analyst plans to explore the effects of different building design options, the distributions for these variables may use uniform distribution. If the purpose is to investigate the possible range of thermal performance for an existing building, the normal distributions may be used for most of variables. More research is needed for the second application.

(5) The variations of sensitivity index need to be computed in order to provide more robust results for sensitivity analysis. Bootstrapping technique can be very useful for this purpose.

(6) Care should be taken when implementing sensitivity analysis in the presence of correlated inputs. This may occur in dealing with the data from building stock research.

## References


[1] Hygh JS, DeCarolis JF, Hill DB, Ranji Ranjithan S. Multivariate regression as an energy assessment tool in early building design. Building and Environment 2012;57:165–75.

[2] Hopfe CJ, Hensen JLM. Uncertainty analysis in building performance simulation for design support. Energy and Buildings 2011;43:2798–805.

[3] Westphal FS, Lamberts R Building simulation calibration using sensitivity analysis, August 15–18, 2005, Montréal, Canada. IBPSA (International Building Performance Simulation Association) conference2005.

[4] Sun J, Reddy TA. Calibration of building energy simulation programs using the analytic optimization approach (RP-1051). HVAC&R Research 2006;12:177–96.

[5] Gustafsson S-I. Sensitivity analysis of building energy retrofits. Applied Energy 1998;61:13–23.

[6] Lam JC, Wan KKW, Yang L. Sensitivity analysis and energy conservation measures implications. Energy Conversion and Management 2008;49:3170–7.

[7] Tian W, Choudhary R. A probabilistic energy model for non-domestic building sectors applied to analysis of school buildings in greater London. Energy and Buildings 2012;54:1–11.

[8] Moran F, Natarajan S, Nikolopoulou M. Developing a database of energy use for historic dwellings in Bath, UK. Energy and Buildings 2012 in press; http://dx.doi.org/10.1016/j.enbuild.2012.09.016.

[9] Tian W, de Wilde P. Uncertainty and sensitivity analysis of building performance using probabilistic climate projections: a UK case study. Automation in Construction 2011;20:1096–109.

[10] de Wilde P, Tian W Preliminary application of a methodology for risk assessment of thermal failures in buildings subject to climate change. Strachan, Kelly, Kummert, editors IBPSA (International Building Performance Simulation Association) conference 2009. p. 2077–2084.

[11] Macdonald I, Strachan P. Practical application of uncertainty analysis. Energy and Buildings 2001;33:219–27.

[12] Mechri HE, Capozzoli A, Corrado V. USE of the ANOVA approach for sensitive building energy design. Applied Energy 2010;87:3073–83.

[13] de Wilde P, Tian W. Identification of key factors for uncertainty in the prediction of the thermal performance of an office building under climate change. Building Simulation 2009;2:157–74.

[14] Saltelli A, Ratto M, Tarantola S, Campolongo F. Update 1 of: sensitivity analysis for chemical models. Chemical Reviews 2012;112:PR1–21.

[15] Mara T, Tarantola S. Application of global sensitivity analysis of model output to building thermal simulations. Building Simulation 2008;1:290–302.

[16] Lomas KJ, Eppel H. Sensitivity analysis techniques for building thermal simulation programs. Energy and Buildings 1992;19:21–44.

[17] Lam JC, Hui S. Sensitivity analysis of energy performance of office buildings. Building and Environment 1996;31:27–39.

[18] Rasouli M, Ge G, Simonson CJ, Besant RW. Uncertainties in energy and economic performance of HVAC systems and energy recovery ventilators due to uncertainties in building and HVAC parameters. Applied Thermal Engineering 2013;50:732–42.

[19] Cammarata G, Fichera A, Marletta L. Sensitivity analysis for room thermal response. International Journal of Energy Research 1993;17:709–18.

[20] Jensen SØ. Validation of building energy simulation programs: a methodology. Energy and Buildings 1995;22:133–44.

[21] Petersen S, Svendsen S. Method and simulation program informed decisions in the early stages of building design. Energy and Buildings 2010;42:1113–9.

[22] Demanuele C, Tweddell T, Davies M Bridging the gap between predicted and actual energy performance in schools. World Renewable Energy Congress XI, 25-30 September 2010, Abu Dhabi, UAE2010.

[23] Firth SK, Lomas KJ, Wright AJ. Targeting household energy-efficiency measures using sensitivity analysis. Building Research & Information 2009;38:25–41.

[24] Cheng V, Steemers K. Modelling domestic energy consumption at district scale: a tool to support national and local energy policies. Environmental Modelling and Software 2011;26:1186–98.

[25] Yildiz Y, Korkmaz K, Göksal Özbalta T, Durmus Arsan Z. An approach for developing sensitive design parameter guidelines to reduce the energy requirements of low-rise apartment buildings. Applied Energy 2012;93:337–47.

[26] Domínguez-Muñoz F, Cejudo-López JM, Carrillo-Andrés A. Uncertainty in peak cooling load calculations. Energy and Buildings 2010;42:1010–8.

[27] Breesch H, Janssens A. Performance evaluation of passive cooling in office buildings based on uncertainty and sensitivity analysis. Solar Energy 2010;84:1453–67.

[28] Struck C, Hensen J, Kotek P. On the application of uncertainty and sensitivity analysis with abstract building performance simulation tools. Journal of Building Physics 2009;33:5–27.

[29] Ballarini I, Corrado V. Analysis of the building energy balance to investigate the effect of thermal insulation in summer conditions. Energy and Buildings 2012;52:168–80.

[30] Tian W, de Wilde P Sampling-based sensitivity analysis of thermal performance in domestic buildings under climate change, EG-ICE(European Group for Intelligent Computing in Engineering), 6–8 JUly 2011, Twente University, The Netherlands 2011.

[31] Wright JA, Wang M, Brownlee AEI, Buswell RA Variable convergence in evolutionary optimization and its relationship to sensitivity analysis, First building simulation and optimization conference, 10–11 September 2012, Loughborough, UK 2012.

[32] Brohus H, Heiselberg P, Simonsen A, Sørensen K Uncertainty of energy consumption assessment of domestic buildings. Proceedings of the eleventh International Building Performance Simulation Association (IBPSA) Conference, Glasgow, UK2009. p. 27–30.

[33] Zhao F Agent-based modeling of commercial building stocks for energy policy and demand response analysis, PhD thesis: Georgia Institute of Technology; 2012.

[34] Zhao J, Xin Y, Tong D. Energy consumption quota of public buildings based on statistical analysis. Energy Policy 2012;43:362–70.

[35] Sanchez DG, Lacarrière B, Musy M, Bourges B. Application of sensitivity analysis in building energy simulations: combining first- and second-order elementary effects methods. Energy and Buildings 2012 in press; http://dx.doi.org/10.1016/j.enbuild.2012.08.048.

[36] Heo Y, Choudhary R, Augenbroe GA. Calibration of building energy models for retrofit analysis under uncertainty. Energy and Buildings 2012;47:550–60.

[37] Corrado V, Mechri HE. Uncertainty and sensitivity analysis for building energy rating. Journal of Building Physics 2009;33:125–56.

[38] Heiselberg P, Brohus H, Hesselholt A, Rasmussen H, Seinre E, Thomas S. Application of sensitivity analysis in design of sustainable buildings. Renewable Energy 2009;34:2030–6.
[39] de Wit S, Augenbroe G. Analysis of uncertainty in building design evaluations and its implications. Energy and Buildings 2002;34:951–8.
[40] Moon HJ. Uncertainty analysis in mould spore transportation and its application in an existing building. Indoor and Built Environment 2010;19:355–65.
[41] Hyun S, Park C, Augenbroe G. Analysis of uncertainty in natural ventilation predictions of high-rise apartment buildings. Building Services Engineering Research and Technology 2008;29:311–26.
[42] Shen H, Tzempelikos A. Sensitivity analysis on daylighting and energy performance of perimeter offices with automated shading. Building and Environment 2012 in press; http://dx.doi.org/10.1016/j.buildenv.2012.08.028.
[43] Ruiz Flores R, Bertagnolio S, Lemort V Global sensitivity analysis applied to total energy use in buildings, international high performance buildings conference, July 16–19, 2012. Purdue.
[44] Spitz C, Mora L, Wurtz E, Jay A. Practical application of uncertainty analysis and sensitivity analysis on an experimental house. Energy and Buildings 2012;55:459–70.
[45] de Wilde P, Tian W. Predicting the performance of an office under climate change: a study of metrics, sensitivity and zonal resolution. Energy and Buildings 2010;42:1674–84.
[46] Eisenhower B, O'Neill Z, Fonoberov VA, Mezić I. Uncertainty and sensitivity decomposition of building energy models. Journal of Building Performance Simulation 2012;5:171–84.
[47] Helton JC, Johnson JD, Sallaberry C, Storlie CB. Survey of sampling-based methods for uncertainty and sensitivity analysis. Reliability Engineering & System Safety 2006;91:1175–209.
[48] Domínguez-Muñoz F, Anderson B, Cejudo-López JM, Carrillo-Andrés A. Uncertainty in the thermal conductivity of insulation materials. Energy and Buildings 2010;42:2159–68.
[49] Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, et al. Global sensitivity analysis: the primer. Wiley-Interscience; 2008.
[50] Tian W, de Wilde P. Thermal building simulation using the UKCP09 probabilistic climate projections. Journal of Building Performance Simulation 2011;4:105–24.
[51] DOE. EnergyPlus V7.1, May 2012, Department of Energy, USA. 2012.
[52] Storlie CB, Swiler LP, Helton JC, Sallaberry CJ. Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models. Reliability Engineering & System Safety 2009;94:1735–63.
[53] Burhenne S, Jacob D, Henze GP Uncertainty analysis in building simulation with Monte Carlo techniques. Fourth national conference of IBPSA-USA, New York, August 11–132010.
[54] Morris MD. Factorial sampling plans for preliminary computational experiments. Technometrics 1991;33:161–74.
[55] Gramacy RB, Taddy M. Categorical inputs, sensitivity analysis, optimization and importance tempering with tgp version 2, an *R* package for treed Gaussian process models. Journal of Statistical Software 2012;33:1–48.
[56] Saltelli A, Tarantola S, Campolongo F, Ratto M. Sensitivity analysis in practice: a guide to assessing scientific models. Wiley; 2004.
[57] Saltelli A. Sensitivity analysis: could better methods be used. Journal of Geophysical Research 1999;104:3789–93.
[58] Hamby D. A review of techniques for parameter sensitivity analysis of environmental models. Environmental Monitoring and Assessment 1994;32:135–54.
[59] SIMLAB. V2.2, Simulation environment for uncertainty and sensitivity analysis. Developed by the Joint Research Center of the European Commission. 2011.
[60] Levy S, Steinberg DM. Computer experiments: a review. AStA Advances in Statistical Analysis 2010;94:311–24.
[61] Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference, and prediction. Springer; 2009.
[62] Oakley JE, O'Hagan A. Probabilistic sensitivity analysis of complex models: a Bayesian approach. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 2004;66:751–69.
[63] Kleijnen JPC. An overview of the design and analysis of simulation experiments for sensitivity analysis. European Journal of Operational Research 2005;164:287–300.
[64] Jolliffe IT. Principal component analysis. London: Wiley; 2002.
[65] Bi J. A review of statistical methods for determination of relative importance of correlated predictors and identification of drivers of consumer liking. Journal of Sensory Studies 2012;27:87–101.
[66] Grömping U. Estimators of relative importance in linear regression based on variance decomposition. The American Statistician 2007;61:139–47.
[67] Grömping U. Variable importance assessment in regression: linear regression versus random forest. The American Statistician 2009;63:308–19.
[68] Strobl C, Boulesteix AL, Kneib T, Augustin T, Zeileis A. Conditional variable importance for random forests. BMC Bioinformatics 2008;9:307.
[69] Efron B, Tibshirani R. An introduction to the bootstrap. Chapman & Hall/CRC; 1993.
[70] Gilles Pujol BI, Alexandre Janon. *R* package sensitivity: sensitivity analysis. ⟨http://cran.r-project.org/package=sensitivity/⟩. 2012.
[71] Grömping U. Relative importance for linear regression in *R*: the package relaimpo. Journal of Statistical Software 2006;17:1–27.
[72] *R* development core team. *R*: a language and environment for statistical computing. *R* foundation for statistical computing, Vienna, Austria. ISBN 3-900051-07-0, URL ⟨http://www.R-project.org/⟩. 2011.
[73] Storlie C CompModSA: sensitivity analysis for complex computer models. ⟨http://cran.r-project.org/package=CompModSA/⟩. 2012.
[74] Torsten Hothorn KH, Carolin Strobl, Achim Zeileis sensitivity: sensitivity analysis. http://cran.r-project.org/package=party/. 2012.